DOCUMENT RESUME

ED 038 404 24 TE 001 783

AUTHOR	Rott, Robert K.; And Others
TITLE	Student Abilities in the Evaluation of Verbal Argument: A Normative Study. Report from Concepts in Verbal Argument Project.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
REPORT NO	TR-106
BUREAU NO	BR-5-0216
PUB DATE	Nov 69
CONTRACT	OEC-5-10-154
NOTE	28p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.50
DESCRIPTORS	*Cognitive Ability, *Cognitive Tests, *Critical Thinking, Grade 10, Intelligence Tests, Language Ability, Listening Comprehension, Logical Thinking, Measurement Instruments, Persuasive Discourse, Reading Tests, Secondary Education, Sex Differences, *Speech Instruction, Test Results, *Verbal Development
IDENTIFIERS	*Wisconsin Tests Testimony Reasoning Assessment, WISTTRA

ABSTRACT

A study conducted in four Wisconsin secondary schools attempted to measure student abilities in evaluating verbal argument and to determine the grade level at which verbal argument concepts and skills should be appropriately taught. Seven tests, known as the Wisconsin Tests of Testimony and Reasoning Assessment (WISTTRA), were administered to 3000 students in grades 7-12; three of the tests measured the ability to detect testimony violating common internal and external tests, and four tests measured the ability to recognize and question essential parts of an argument and to draw appropriate conclusions from it. Data indicated that (1) since the greatest change in mean scores between adjacent grades occurred between grades 9 and 10, grade 10 may be the optimum time to teach these abilities, (2) females seem to acquire these critical abilities earlier than do males, and (3) student scores on verbal argument correlated only low to moderate with their scores on intelligence and reading tests. (Included are charts which correlate test data at each grade level.) (See also ED 036 521, TE 001 784, and ED 016 658.) (JB)

Technical Report No. 106

STUDENT ABILITIES IN THE EVALUATION OF VERBAL ARGUMENT: A NORMATIVE STUDY

By Robert K. Rott, Jerry D. Feezel,
R. R. Allen, and Margaret L. Harris

Report from Concepts in Verbal Argument Project
R. R. Allen, Principal Investigator

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

November 1969

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

NATIONAL EVALUATION COMMITTEE

Samuel Brownell
Professor of Urban Education
Graduate School
Yale University

Henry Chauncey
President
Educational Testing Service

Elizabeth Koontz
President
National Education Association

Patrick Suppes
Professor
Department of Mathematics
Stanford University

Launor F. Carter
Senior Vice President on Technology and Development
System Development Corporation

Martin Deutsch
Director, Institute for Developmental Studies
New York Medical College

Roderick McPhee
President
Punahou School, Honolulu

***Benton J. Underwood**
Professor
Department of Psychology
Northwestern University

Francis S. Chase
Professor
Department of Education
University of Chicago

Jack Edling
Director, Teaching Research Division
Oregon State System of Higher Education

G. Wesley Sowards
Director, Elementary Education
Florida State University

UNIVERSITY POLICY REVIEW BOARD

Leonard Berkowitz
Chairman
Department of Psychology

John Guy Fowlkes
Director
Wisconsin Improvement Program

Herbert J. Klausmeier
Director, R & D Center
Professor of Educationol Psychology

M. Crawford Young
Associate Dean
The Graduate School

Archie A. Buchmiller
Deputy State Superintendent
Department of Public Instruction

Robert E. Grinder
Chairman
Department of Educational Psychology

Donald J. McCarty
Dean
School of Education

***James W. Cleary**
Vice Chancellor for Academic Affairs

H. Clifton Hutchins
Chairman
Department of Curriculum and Instruction

Ira Sharkansky
Associate Professor of Political Science

Leon D. Epstein
Dean
College of Letters and Science

Clauston Jenkins
Assistant Director
Coordinating Committee for Higher Education

Henry C. Weinlick
Executive Secretary
Wisconsin Education Association

EXECUTIVE COMMITTEE

Edgar F. Borgatta
Brittingham Professor of Sociology

Russell J. Hosler
Professor of Curriculum and Instruction and of Business

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Richard L. Venezky
Assistant Professor of English and of Computer Sciences

Max R. Goodson
Professor of Educational Policy Studies

***Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational Psychology

Robert G. Petzold
Associate Dean of the School of Education
Professor of Curriculum and Instruction and of Music

FACULTY OF PRINCIPAL INVESTIGATORS

Ronald R. Allen
Associate Professor of Speech and of Curriculum and Instruction

Gary A. Davis
Associate Professor of Educational Psychology

Max R. Goodson
Professor of Educational Policy Studies

Richard G. Morrow
Assistant Professor of Educational Administration

Vernon L. Allen
Associate Professor of Psychology
(On leave 1968-69)

M. Vere DeVault
Professor of Curriculum and Instruction (Mathematics)

Warren O. Hagstrom
Professor of Sociology

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Nathan S. Blount
Associate Professor of English and of Curriculum and Instruction

Frank H. Farley
Assistant Professor of Educational Psychology

John G. Harvey
Associate Professor of Mathematics and Curriculum and Instruction

Milton O. Pella
Professor of Curriculum and Instruction (Science)

Robert C. Calfee
Associate Professor of Psychology

John Guy Fowlkes (Advisor)
Professor of Educational Administration
Director of the Wisconsin Improvement Program

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

Thomas A. Romberg
Assistant Professor of Mathematics and of Curriculum and Instruction

Robert E. Davidson
Assistant Professor of Educational Psychology

Lester S. Golub
Lecturer in Curriculum and Instruction and in English

Burton W. Kreitlow
Professor of Educational Policy Studies and of Agricultural and Extension Education

Richard L. Venezky
Assistant Professor of English and of Computer Sciences

MANAGEMENT COUNCIL

***Herbert J. Klausmeier**
Director, R & D Center
Acting Director, Program 1

Thomas A. Romberg
Director
Programs 2 and 3

James E. Walter
Director
Dissemination Section

Dan G. Woolpert
Director
Operations and Business

Mary R. Quilling
Director
Technical Section

* COMMITTEE CHAIRMAN

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Concepts in Verbal Argument Project in Program 2. General objectives of the Program are: to establish rationale and strategy for developing instructional systems, to identify sequences of concepts and cognitive skills, to develop assessment procedures for those concepts and skills, to identify or develop instructional materials associated with the concepts and cognitive skills, and to generate new knowledge about instructional procedures. Contributing to these Program objectives, the staff of the project developed a semiprogramed course in verbal argument and related tests for use at the high school level. The project staff prepared the materials on the basis of an outline of concepts and critical skills developed from an evaluation of everyday discourse.

CONTENTS

LIST OF TABLES
In Text

LIST OF TABLES
In Appendix

ABSTRACT

The normative study reported herein was conducted during the 1968 spring semester in the junior and senior high schools of Clinton, Cedarburg, Reedsburg, and Owen-Withee, Wisconsin. A battery of seven tests, known collectively as the Wisconsin Tests of Testimony and Reasoning Assessment (WISTTRA), was administered to more than 3000 participating subjects.

Inspection of the data seems to warrant the following conclusions: mean scores tend to increase gradually from Grade 7 through Grade 12; (the greatest change in mean scores between adjacent grades occurs between Grades 9 and 10) thus, Grade 10 may be an optimum time to teach the abilities necessary for verbal argument; male students appear to acquire the abilities tested a little later in life than do females; and correlations of student scores on the verbal argument tests with student scores on intelligence and reading tests tend to be low to moderate.

I

INTRODUCTION

The normative study reported in this paper constitutes one phase of the research and development activities conducted by the Concepts in Verbal Argument Project. This phase was preceded by a review of relevant literature (Allen & Rott, 1969), the development of a taxonomy of concepts and critical abilities related to the evaluation of verbal argument (Allen, Feezel, & Kauffeld, 1967), the construction of a battery of tests to measure student abilities in assessing testimony and reasoning (Allen, Feezel, & Kauffeld, 1969), the testing and revision of the measuring instrument (Allen, Feezel, Kauffeld, & Harris, 1969), and a study of the dimensionality of the tests using factor-analytic procedures (Harris, 1969).

Since the terminal goal of the project involves the making available of learning materials in verbal argument for use by high school students, it was decided early in the project that data should be gathered regarding the critical abilities of the general target population. Such data would provide a basis for determining at what grade level instruction in these concepts would seem most appropriate. It was also intended that the normative data would guide the authors in the preparation of the learning program by offering precise information regarding pre-instructional student skills. Since the Wisconsin Tests of Testimony and Reasoning Assessment (WISTTRA) were developed around the particular skills and their underlying conceptual structure of direct interest to the investigators, the data provided by the administration of the battery is highly specific and, therefore, directive. The learning program (Allen, Kauffeld, & O'Brien, 1968), field-tested during the 1969 Spring Semester (Allen, Fischbach, & Quilling, in press), was strongly influenced by the data gathered in the normative study.

In addition to its value to the project at large, the normative study may be found of interest in and of itself. In seeking to ascertain the ability of sutdents, Grades 7 through 12, to evaluate testimony and arguments established through reasoning, this study provides data of interest to critical thinking researchers whatever their conceptual bias and to high school teachers in several academic disciplines.

II

PROCEDURES

In this section may be found a discussion of the measurement instrument for which norms were developed, the population from which the sample of subjects was drawn, administration procedures, and the method of analyzing and presenting the data.

MEASUREMENT INSTRUMENT

The Wisconsin Tests of Testimony and Reasoning Assessment (WISTTRA) were developed by the researchers of the Concepts in Verbal Argument Project at the Wisconsin Research and Development Center for Cognitive Learning. WISTTRA was developed to assess the student's ability to evaluate adequacy of testimony, to recognize the structure that is present in ordinary arguments, and to raise pertinent objections based on the rules of inference appropriate to that structure. There are seven tests in the WISTTRA battery. The three tests of testimony were designed to measure the ability to detect instances of testimony which violate common internal and external tests. The four tests of reasoning were designed to measure the ability to recognize the essential parts of an argument, to ask relevant questions about arguments, and to draw appropriate conclusions from arguments. A thorough discussion of the development of the tests, including a statement of the objective for each test, a brief description of each test, and reliability estimates and summary item statistics for each test, can be found in Allen, Feezel, Kauffeld, and Harris (1969). The tests proper are contained in Allen, Feezel, and Kauffeld (1969). The names of the tests are: Testimony I consisting of 1 Accept subtest and 4 Reject subtests—Bias, Position, Competence, and Qualification; Testimony II; Testimony III consisting of 2 subtests—Recency and Proximity; Reasoning I; Reasoning II; Reasoning III; and Reasoning IV.[1]

[1]See footnote on page 27.

From two of the schools for which they were available, Cedarburg and Reedsburg, intelligence and reading scores were secured. The nature of these scores and the tests which were used to obtain them are as follows:

Cedarburg:

Grade 7: Intelligence scores are nonverbal raw scores from Lorge-Thorndike Form 1, Level E.
Reading scores are national percentiles from Iowa Tests of Basic Skills, Form 2.

Grade 8: Intelligence scores are verbal raw scores from Lorge-Thorndike Form AA, Level 4.
Reading scores are the same as for Grade 7.

Grades 9-12: Intelligence scores are raw scores from Henmon-Nelson Test of Mental Maturity.
Reading scores are standard scores from Iowa Tests of Educational Development, Reading Average Section, Form XP-CP.

Reedsburg:

Grades 7-8: Intelligence scores are I.Q. scores from the Henmon-Nelson Test of Mental Ability.
Reading scores are converted grade-placement scores from the Iowa Tests of Basic Skills.

Grade 9: Intelligence scores are raw scores from the Henmon-Nelson Test of Mental Ability.
Reading scores are raw scores from the Iowa Tests of Educational Development.

Grades 10-12: Intelligence scores are Henmon-Nelson raw scores.
Reading scores are ITED raw scores.

SUBJECTS

The researchers wished to obtain norms for the WISTTRA battery for boys and girls separately in Grades 7 through 12. The formula

$$\left[z_{\frac{(1-\alpha)}{2}}\right]\frac{\sigma}{\sqrt{N}} = \text{tolerance limit}$$

was used to determine the size sample necessary to obtain a sample (estimated) mean that would be within a certain tolerance limit of the population (true) mean with a certain level of confidence. Using a standard deviation of seven (which is an approximate upper size for the tests being normed) and an alpha of .05, it was found that a sample size of approximately 200 would be necessary for a tolerance limit of one unit. This is the sample size necessary for each sex and grade group since the data will be analyzed separately for each of the twelve groups.

To obtain a homogeneous population from which to sample, information obtained in a survey of Wisconsin school districts (Miller, et al., 1967) was used. This survey placed 674 Wisconsin school districts into 32 strata on the basis of their factor scores on the following six factors:

1. Numerical size
2. Organizational complexity
3. Teacher experience
4. Economic power
5. Size of school unit
6. Log variance of teacher credentials

For stratification purposes, a school district was given a + or - for each of the six factors; a plus signified a factor score at or above the median and a minus a factor score below the median.

The school districts from two of these strata were used as the population for this normative study. These two strata were, + + + - - + and + + + - - -. Thus the schools comprising the population were above the median in numerical size, organizational complexity, and teacher experience. A plus for Factor 4, economic power, means that the district has a high valuation and such districts tend to be elementary only school districts; a minus was used for this factor. A plus for Factor 5, school unit size, means that the district has a large number of one-and two-room schools; a minus was used for this factor also. Since two of the strata were used, the one having a plus for Factor 6 as well as the one having a minus for Factor 6, log variance of teacher credentials was not a factor characteristic of the population used.

A total of 19 schools was included in the population. Of these, 13 are city and 6 are county school districts. Eighteen have Grades K-12; the one exception, which was not one of the sampled schools, has Grades 1-12. All are receiving state aid. On the average their secondary enrollment is 1366.2 students, size of staff is 144.3, and on the average these districts have no one-room schools, .15 two-room schools, and 7.1 schools of three or more rooms. Also the following statistics, by strata, may be of interest:

	+ + + - - -	+ + + - - +
Valuation per school	$12,349,427.586	$10,563,546.789
Valuation per student	$27,259.355	$23,358.352
Students per school	453.034	452.239
Students per staff	22.969	22.727
Staff per school	19.724	19.899

A random sample of six schools was drawn from the 19 schools comprising the population. These schools were then contacted, in the order in which they were sampled, and asked to participate in the study. Three of them agreed to do so. The total number of students in Grade 12 (Grade 12 was assumed to have the fewest and was used as the base) was not as many as the 400 we wanted—200 boys and 200 girls. Further random sampling was conducted and the school districts were contacted. The first additional district sampled was unable to participate but the second one did take part in the study. Thus the subjects for this study were from four school districts in the state of Wisconsin—Clinton, Cedarburg, Reedsburg, and Owen-Withee—with the characteristics previously explained. The size of the total sample was 3090 subjects to 3105 subjects depending upon the particular test. The smallest group in the sample was 190 Twelfth Grade males; all other groups contained well over 200

subjects. The exact number of subjects per group for each test can be found in Tables A through O, in the Appendix.

ADMINISTRATION OF WISTTRA BATTERY

The test battery consisting of the three testimony tests and four reasoning tests was administered during successive weeks in April 1968 to students in Grades 7 through 12 in the following Wisconsin junior and senior high schools in the order indicated: Clinton, Cedarburg, Reedsburg, and Owen-Withee. Each student was given as much time as needed to complete the tests. To minimize the fatigue factor, testing was conducted in two separate sessions on successive days, afternoon session followed by morning session. Testing time required for each session was from 2 to 3 hours. There was great variation in the time required by students to complete the tests; however, almost all students were able to finish either of the test sessions in less than 3 hours. It is estimated that the average completion time for the majority of students was about 2 1/2 hours, with a substantial number finishing in about 2 hours and a few finishing in less than 2 hours.

The order of test taking was randomized, within certain tests allocated for a single session, across the four schools to cancel possible learning effects of tests taken in their logical order. Table 1 includes the testing orders for the testimony (T) and reasoning (R) tests for the two testing sessions in the four schools.

The students responded to the tests by marking their chosen answer directly on an answer sheet. Each student was given an envelope containing the printed test booklets and answer sheets for each session. The envelope was marked with the prescribed order in which the tests were to be taken. The test booklets were also arranged in this order.

TREATMENT OF DATA

Hoyt analysis of variance reliability estimates were obtained for each of the total tests and for subtests of Testimony I and Testimony III for each sex and grade group. The standard error of measurement was also computed for each of these.

A summary of the item statistics for each test can be found in Allen, Feezel, Kauffeld, and Harris (1969).

The mean and standard deviation were computed for each sex and grade group for each total test and for the subtests of Testimony I and Testimony III. The difference between the means of adjacent grades was found, by sex, for each of the total tests.

Intercorrelations of the seven tests in the WISTTRA battery were obtained. Included also were the intercorrelations of these seven tests with Testimony I as two subtests and with Testimony III as two subtests.

Intercorrelations of the seven tests in the WISTTRA battery with intelligence and reading scores were obtained for subjects from two of the schools, Cedarburg and Reedsburg.

The next section will present the results obtained and a discussion of these results.

Table 1

Testing Orders

School	First Day	Second Day
1. Clinton	R-IV, R-I, R-III	T-II, T-I, T-III, R-II
2. Cedarburg	R-III, R-I, R-IV	T-I, T-II, T-III, R-II
3. Reedsbrug	T-III, R-II, T-I, T-II	R-IV, R-III, R-I
4. Owen-Withee	T-II, T-I, R-II, T-III	R-I, R-IV, R-III

III
RESULTS AND DISCUSSION

In this section information will be presented regarding reliability estimates, differences of means of adjacent grades, intercorrelations of tests which comprise the WISTTRA battery, and correlations of WISTTRA with intelligence and reading. The reader is reminded that a description of the tests can be found in Allen, Feezel, Kauffeld, and Harris (1969).

RELIABILITY ESTIMATES

Tables A through O in the Appendix contain the reliability estimates, standard errors of measurement, means, and standard deviations for each sex and grade group for each of the total tests, for each of the subtests of Testimony I and Testimony III, and for the four reject subtests of Testimony I as a unit.

The reliability estimates for the total tests, in general, gradually increase in magnitude from Grade 7 through Grade 12. For Grades 7 through 9 the estimates tend to be in the .70s to low .80s. For Grades 10 through 12 they tend to be in the .80s to low .90s. These reliability estimates are adequate for research purposes and for assessing level of group achievement; some of the reliabilities are adequate for assessing level of individual achievement. The reliability estimates are very similar at a particular grade level for the two sexes.

DIFFERENCES OF MEANS

The mean scores tend to increase gradually from Grade 7 through Grade 12 with the standard deviation remaining fairly similar in most cases. Table 2 contains the differences, by sex, between the mean scores for adjacent grades. With three exceptions, the magnitude of the differences between the means of adjacent grades

Table 2

Difference Between Means of Adjacent Grades

Grade and Sex	T-I	T-II	T-III	R-I	R-II	R-III	R-IV
7M-8M	1.8864	-0.4671*	0.4502	0.1722	1.8146	1.4503	1.6687
7F—8F	1.9103	0.3918	1.6973	1.8213	1.0121	1.9500	1.5331
8M-9M	1.2160	0.0822	0.3597	0.2625	0.2423	0.7719	0.5428
8F—9F	1.3447	0.1009	0.4243	0.8256	1.3030	0.8479	0.8835
9M-10M	2.6967	0.5742	1.8854	3.2198	0.8819	2.8862	2.1191
9F—10F	2.3179	1.8775	1.3793	2.8616	1.4427	2.3748	1.9074
10M-11M	1.7572	1.2772	1.0434	1.1849	1.9267	1.1381	0.5144
10F—11F	-0.1482*	-0.3127*	1.3041	0.7349	0.4994	0.3830	0.4354
11M-12M	-0.4717*	-0.0836*	-0.5076*	-0.2216*	0.2119	-0.6570*	-0.1962*
11F—12F	0.9155	0.1578	-0.6735*	0.4913	0.4082	1.0654	0.9960

* A negative sign indicates that as grade increases, the mean score from one grade to the next decreases.

is the greatest for any one test between Grades 9 and 10. For males the magnitude of the difference is the greatest between Grades 10 and 11 for Testimony II and Reasoning II. The one exception for females occurs for Testimony III; the difference is the greatest between Grades 7 and 8. These two exceptions for males and the fact that the differences between Grades 10 and 11 tend to be higher for males than for females are indications that male students may acquire the abilities tested a little later in life than do females. From looking at the total pattern of mean differences between adjacent grades it seems that 10th Grade may be a good time to teach these types of verbal argument skills. During this period students are gaining the most in these skills without instruction and thus, it may be the best time to supplement this natural learning with instruction.

INTERCORRELATIONS OF WISTTRA

The intercorrelations of the various tests included in the WISTTRA battery are given in Tables 3 through 7 for males and females in Grades 7 through 12.

The three testimony tests and four reasoning tests were constructed so that each was composed of at least two subtests. The study of the dimensionality of the tests indicates that the four reject subtests of Testimony I (Bias, Position, Competence, and Qualification) are a separate dimension and that Testimony III consists of two different dimensions, Recency and Proximity (Harris, 1969). Thus, these tables include the intercorrelations of Testimony I; Testimony II; Testimony III; Reasoning I; Reasoning II; Reasoning III; Reasoning IV; Testimony I, Accept; Testimony I, Reject; Testimony III, Recency; and Testimony III, Proximity.

The correlations, in general, are low to moderate indicating that these separate tests in the WISTTRA battery are measuring at least somewhat different abilities. The fairly high correlations that appear are ones that would be expected to be high—for TI-Accept and TI-Reject with TI, and for TIII-Recency and TIII-Proximity with TIII —since one is correlating a part of the test with the total test of which it is a part.

CORRELATIONS OF WISTTRA WITH INTELLIGENCE AND READING

The correlations of the three tests of testimony and the four tests of reasoning with intelligence and reading scores are given in Table 9 for Cedarburg and Table 10 for Reedsburg. Intelligence and reading scores were not available for the other two communities tested. The particular tests used to obtain the intelligence and

Table 3

Intercorrelations of WISTTRA for Grade 7 Males and Females*

	TI		TII		TIII		RI		RII		RIII		RIV		TI-A		TI-R		TIII-R	
Sex	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
TII	48	43																		
TIII	47	48	46	52																
RI	38	61	43	49	53	60														
RII	49	55	49	49	50	59	58	59												
RIII	60	53	54	54	60	59	51	68	66	66										
RIV	48	52	48	48	52	49	56	55	60	64	63	62								
TI-A	66	58	52	44	45	49	33	48	47	48	64	52	51	49						
TI-R	75	78	21	20	25	22	22	38	27	29	26	25	23	26	06	01				
TIII-R	42	47	41	48	86	83	47	53	47	53	56	57	47	43	39	49	22	21		
TIII-P	39	30	36	36	86	80	43	45	39	42	48	39	42	36	38	29	21	15	46	33

* Decimals have been omitted.

Table 4

Intercorrelations of WISTTRA for Grade 8 Males and Females*

	TI		TII		TIII		RI		RII		RIII		RIV		TI-A		TI-R		TIII-R	
Sex	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
TII	54	40																		
TIII	63	50	52	32																
RI	57	51	59	49	57	48														
RII	61	50	58	48	52	48	65	61												
RIII	69	59	53	52	64	53	69	63	68	68										
RIV	63	56	53	46	55	48	60	60	58	72	70	77								
TI-A	78	70	52	38	59	45	53	53	51	53	66	62	62	60						
TI-R	84	80	36	21	43	30	38	27	45	22	47	32	42	28	37	20				
TIII-R	59	47	47	33	84	82	55	52	48	49	61	55	55	48	51	44	45	28		
TIII-P	46	32	40	17	83	76	41	23	39	26	46	28	36	27	47	27	27	19	40	26

* Decimals have been omitted.

Table 5

Intercorrelations of WISTTRA for Grade 9 Males and Females*

	TI		TII		TIII		RI		RII		RIII		RIV		TI-A		TI-R		TIII-R	
Sex	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
TII	45	41																		
TIII	44	50	42	43																
RI	54	58	54	58	48	50														
RII	46	48	51	60	46	43	64	66												
RIII	62	57	55	57	50	42	64	68	65	70										
RIV	54	57	52	53	40	47	62	66	63	68	69	73								
TI-A	74	56	50	43	46	39	56	49	49	45	64	58	58	51						
TI-R	85	83	28	21	28	31	31	36	27	28	37	31	32	34	32	06				
TIII-R	44	47	41	44	86	85	49	50	43	40	46	44	41	44	45	37	27	31		
TIII-P	31	38	30	30	84	85	32	35	35	33	38	28	27	36	33	30	19	22	45	44

* Decimals have been omitted.

Table 6

Intercorrelations of WISTTRA for Grade 10 Males and Females*

	TI		TII		TIII		RI		RII		RIII		RIV		TI-A		TI-R		TIII-R	
Sex	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
TII	53	46																		
TIII	58	40	56	44																
RI	59	55	51	69	69	47														
RII	51	49	56	65	58	44	63	71												
RIII	57	62	49	66	52	47	59	73	71	74										
RIV	53	56	44	61	50	47	65	74	67	67	73	73								
TI-A	75	56	51	44	58	34	52	49	47	44	54	56	45	52						
TI-R	85	87	36	28	38	28	43	40	35	35	38	43	40	39	32	14				
TIII-R	57	40	46	42	86	80	67	46	54	47	51	50	50	47	52	32	41	28		
TIII-P	44	26	51	31	87	84	52	32	48	27	40	28	36	32	49	24	24	19	50	35

* Decimals have been omitted.

Table 7

Intercorrelations of WISTTRA for Grade 11 Males and Females*

	TI		TII		TIII		RI		RII		RIII		RIV		TI-A		TI-R		TIII-R	
Sex	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
TII	33	38																		
TIII	45	36	49	42																
RI	55	53	55	57	63	50														
RII	42	49	62	62	58	47	70	72												
RIII	50	57	60	55	51	49	67	75	71	75										
RIV	46	52	55	53	55	42	67	73	69	68	69	76								
TI-A	66	44	39	46	46	40	55	57	43	48	54	61	50	58						
TI-R	87	85	16	15	29	18	34	28	26	27	27	30	26	26	25	-03				
TIII-R	44	40	49	41	87	84	60	50	55	45	48	49	52	46	44	39	30	22		
TIII-P	34	24	37	32	88	88	50	36	47	37	42	36	44	28	37	31	21	10	53	48

* Decimals have been omitted.

Table 8

Intercorrelations of WISTTRA for Grade 12 Males and Females*

	TI		TII		TIII		RI		RII		RIII		RIV		TI-A		TI-R		TIII-R	
Sex	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
TII	51	29																		
TIII	58	39	57	38																
RI	61	50	60	49	68	51														
RII	59	45	67	54	66	43	71	69												
RIII	61	57	65	51	64	42	71	72	76	71										
RIV	62	55	60	46	62	49	69	68	75	69	83	71								
TI-A	74	43	53	40	60	45	56	51	58	43	63	60	65	58						
TI-R	87	88	34	13	38	20	45	29	41	28	38	34	40	32	35	02				
TIII-R	54	38	57	36	88	84	60	48	58	39	59	42	57	44	61	42	32	22		
TIII-P	48	30	43	31	88	90	60	41	58	37	54	33	53	41	46	38	35	13	55	52

* Decimals have been omitted.

Table 9

Intercorrelations of Intelligence and Reading with Testimony and Reasoning for Grades 7-12, Cedarburg

	Int.	Rdg.	T-I	T-II	T-III	R-I	R-II	R-III	R-IV		N (range)
Grade 7M	1.000	.668	.583	.500	.337	.418	.470	.600	.550	Intelligence	65-79
		1.000	.672	.655	.579	.451	.521	.759	.658	Reading	
7F	1.000	.494	.465	.444	.342	.549	.472	.560	.574	Intelligence	64-77
		1.000	.585	.677	.532	.688	.726	.799	.694	Reading	
Grade 8M	1.000	.808	.655	.497	.558	.695	.707	.704	.754	Intelligence	51-59
		1.000	.771	.591	.624	.677	.702	.786	.792	Reading	
8F	1.000	.791	.451	.624	.507	.757	.708	.810	.808	Intelligence	53-65
		1.000	.502	.504	.494	.609	.695	.747	.651	Reading	
Grade 9M	1.000	.749	.683	.550	.416	.476	.500	.648	.604	Intelligence	96-98
		1.000	.613	.571	.502	.579	.628	.669	.689	Reading	
9F	1.000	.735	.587	.553	.561	.571	.614	.684	.723	Intelligence	87-89
		1.000	.556	.634	.536	.666	.731	.720	.772	Reading	
Grade 10M	1.000	.670	.568	.569	.549	.585	.550	.530	.483	Intelligence	90-91
		1.000	.496	.678	.661	.727	.672	.695	.679	Reading	
10F	1.000	.785	.522	.580	.450	.688	.683	.710	.656	Intelligence	95-98
		1.000	.508	.653	.363	.718	.668	.692	.664	Reading	
Grade 11M	1.000	.835	.413	.610	.548	.582	.632	.611	.599	Intelligence	85-87
		1.000	.521	.658	.647	.699	.726	.686	.670	Reading	
11F	1.000	.824	.547	.553	.482	.737	.667	.719	.640	Intelligence	78-78
		1.000	.517	.560	.597	.750	.745	.732	.749	Reading	
Grade 12M	1.000	.665	.417	.410	.510	.562	.469	.683	.566	Intelligence	57-61
		1.000	.423	.474	.562	.604	.524	.619	.609	Reading	
12F	1.000	.843	.463	.406	.426	.581	.624	.613	.643	Intelligence	90-93
		1.000	.495	.517	.455	.654	.682	.645	.735	Reading	

Table 10

Intercorrelations of Intelligence and Reading with Testimony and Reasoning for Grades 7-12, Reedsburg

	Int.	Rdg.	T-I	T-II	T-III	R-I	R-II	R-III	R-IV		N(range)
Grade 7M	1.000	.774	.586	.459	.398	.554	.591	.573	.562	Intelligence	77-77
		1.000	.664	.550	.427	.724	.567	.714	.685	Reading	
7F	1.000	.663	.344	.299	.333	.359	.386	.598	.488	Intelligence	72-74
		1.000	.512	.391	.415	.310	.493	.641	.655	Reading	
Grade 8M	1.000	.814	.520	.389	.258	.492	.553	.664	.650	Intelligence	75-79
		1.000	.602	.474	.361	.533	.669	.731	.684	Reading	
8F	1.000	.826	.592	.379	.387	.688	.702	.629	.637	Intelligence	66-68
		1.000	.652	.444	.375	.685	.718	.699	.742	Reading	
Grade 9M	1.000	.695	.486	.460	.291	.561	.496	.630	.644	Intelligence	97-99
		1.000	.592	.443	.298	.517	.602	.696	.675	Reading	
9F	1.000	.794	.571	.648	.454	.606	.687	.724	.701	Intelligence	109-113
		1.000	.676	.675	.492	.722	.686	.804	.739	Reading	
Grade 10M	1.000	.798	.565	.338	.344	.526	.560	.667	.618	Intelligence	54-82
		1.000	.670	.355	.366	.582	.607	.754	.661	Reading	
10F	1.000	.862	.520	.481	.343	.603	.711	.720	.656	Intelligence	52-79
		1.000	.469	.504	.327	.595	.650	.663	.625	Reading	
Grade 11M	1.000	.715	.689	.464	.485	.539	.502	635	.570	Intelligence	85-90
		1.000	.653	.504	.572	.608	.630	.727	.677	Reading	
11F	1.000	.732	.516	.513	.420	.616	.624	.617	.658	Intelligence	91-92
		1.000	.580	.551	.406	.660	.714	.679	.701	Reading	
Grade 12M	1.000	.816	.643	.532	.603	.605	.612	.616	.693	Intelligence	66-69
		1.000	.776	.599	.639	.751	.749	.763	.803	Reading	
12F	1.000	.793	.648	.492	.430	.640	.584	.736	.530	Intelligence	74-78
		1.000	.640	.478	.473	.693	.696	.755	.537	Reading	

reading scores are listed in the previous section of this paper. The number of subjects is given for each group, the range being over the seven tests of verbal argument.

The intercorrelations of reading and intelligence are, in most cases, higher than the intercorrelations of reading or intelligence with the testimony and reasoning tests. Correlations between intelligence and reading tend to be fairly high while the correlations of the verbal argument tests with intelligence and reading tend to be low to moderate in magnitude. From this it would appear that the tests in the WISTTRA battery are measuring something different from the abilities measured by the intelligence and reading tests. It should be noted at this point that in the development of the tests, Grade 9 was selected as the maximum difficulty level of the vocabulary used. The difficulty levels indicated in Thorndike and Lorge (1944) were used as a guide.

In comparing the correlations of the testimony tests with intelligence and reading, and the correlations of the reasoning tests with intelligence and reading (comparison by columns), it appears that the tests of reasoning tend to correlate more highly with intelligence and reading than do the tests of testimony, except for Seventh Grade males at Cedarburg. This higher correlation of the reasoning tests with the intelligence measures may indicate that the evaluation of reasoning is more dependent upon an individual's basic mental capacity than are the skills for assessing instances of testimony. The differences in the correlations with reading ability may suggest a similar dependence upon reading skills or may be merely a function of the longer test items with several interrelated sentences which compose the reasoning tests.

In comparing the correlations of the testimony and reasoning tests with intelligence, and the correlations of the testimony and reasoning tests with reading (comparison by rows), it appears that for any one grade level the testimony and reasoning tests tend to correlate more highly with reading than with intelligence, except for the Eighth Grade females at Cedarburg

and the Tenth Grade females at Reedsburg. Perhaps this trend reflects that the evaluation of arguments couched in ordinary language relies upon the analysis of that language and is, therefore, more related to verbal abilities than to a general intellectual ability. Supportive of this hypothesis is that for one of the exceptions to the trend, Eighth Grade females at Cedarburg for which the tests are more highly correlated with intelligence than with reading, the intelligence measure was the verbal portion only from a standard intelligence test.

IV

CONCLUSIONS

The reliability estimates for each of the seven tests in the WISTTRA battery are fairly high. They are substantially higher for Grades 10 through 12, the levels for which the tests were primarily designed, than they are for Grades 7 through 9. The reliability estimates are high enough to warrant assessing levels of group achievement for all grades, 7 through 12. The reliability estimates for Grades 10 through 12 approach desirable levels for assessing individual achievement. Care should be exercised in comparing the scores of individual students when the reliability estimate is below .90; thus, this should not be done for Grades 7 through 9 and done with caution for most of the tests for Grades 10 through 12.

The mean scores tend to increase gradually from Grade 7 through Grade 12. The greatest change in mean scores between adjacent grades occurred, almost always, between Grades 9 and 10. Since the tests were administered in April for this normative study, it would seem that the factors, whatever they may be, that are responsible for the changes in verbal argument abilities are operating with maximum effect after Grade 9 and during Grade 10. Grade 10, therefore, may be a good time to teach these verbal argument abilities.

Males tend to acquire the abilities tested a little later in life than do females.

The intercorrelations of the WISTTRA battery, in general, are low to moderate indicating that these separate tests are measuring at least somewhat different abilities. It appears that the tests in the WISTTRA battery are measuring something different from the abilities measured by the intelligence and reading tests.

APPENDIX

Table A

Testimony I: Reliability Estimates for the Total Test

Grade/Sex	N	Mean Score (60 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	37.89	7.44	18–58	3.48	.78
7F	251	38.96	6.62	12–56	3.42	.73
8M	228	39.78	7.85	24–54	3.38	.81
8F	224	40.87	7.06	25–55	3.29	.78
9M	304	41.00	8.04	13–56	3.29	.83
9F	302	42.22	7.23	25–57	3.19	.80
10M	287	43.69	8.04	6–58	3.08	.85
10F	302	44.53	6.70	23–57	2.97	.80
11M	253	45.45	6.84	23–57	2.92	.81
11F	265	44.38	6.45	16–56	2.96	.79
12M	190	44.98	8.09	22–59	2.97	.86
12F	253	45.30	6.73	12–57	2.88	.81

Table B

Testimony I: Reliability Estimates for the Accept Subtest

Grade/Sex	N	Mean Score (20 max.)	Standard Deviation	Spread of Scores	Standard Error or Measurement	Hoyt Reliability
7M	246	14.22	3.90	0–20	1.80	.78
7F	251	15.00	3.23	5–20	1.75	.69
8M	228	15.00	3.50	7–20	1.74	.74
8F	224	15.65	3.16	7–20	1.64	.72
9M	304	15.42	3.57	6–20	1.66	.77
9F	302	16.19	3.05	5–20	1.54	.73
10M	287	16.43	3.46	7–20	1.47	.80
10F	302	17.14	2.48	8–20	1.37	.68
11M	253	17.50	2.61	7–20	1.31	.74
11F	265	17.38	2.53	6–20	1.31	.72
12M	190	17.14	3.21	7–20	1.37	.81
12F	253	17.68	2.41	4–20	1.23	.72

Table C

Testimony I: Reliability Estimates for the Bias Subtest

Grade/Sex	N	Mean Score (10 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	5.16	2.17	0–9	1.32	.59
7F	251	5.19	2.14	1–10	1.29	.60
8M	228	5.48	2.05	0–10	1.34	.53
8F	224	5.38	2.28	0–10	1.26	.66
9M	304	5.53	2.15	0–10	1.33	.57
9F	302	5.49	2.12	0–10	1.27	.60
10M	287	5.86	2.13	0–10	1.27	.61
10F	302	5.57	2.20	1–10	1.21	.66
11M	253	6.00	2.21	1–10	1.21	.67
11F	265	5.62	2.19	0–10	1.20	.67
12M	190	5.91	2.15	0–10	1.26	.61
12F	253	5.74	2.31	0–10	1.19	.70

Table D

Testimony I: Reliability Estimates for the Position Subtest

Grade/Sex	N	Mean Score (10 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	5.65	1.91	0–10	1.41	.39
7F	251	5.59	2.00	1–10	1.40	.46
8M	228	5.91	1.92	0–10	1.40	.41
8F	224	5.91	1.90	1–10	1.37	.43
9M	304	6.19	1.88	2–10	1.36	.42
9F	302	6.36	2.09	0–10	1.32	.56
10M	287	6.50	1.96	1–10	1.30	.51
10F	302	6.66	1.90	1–10	1.27	.50
11M	253	6.47	1.92	1–10	1.29	.50
11F	265	6.42	1.93	2–10	1.28	.51
12M	190	6.54	2.08	1–10	1.26	.59
12F	253	6.67	2.03	1–10	1.22	.60

Table E

Testimony I: Reliability Estimates for the Competence Subtest

Grade/Sex	N	Mean Score (10 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	6.55	2.03	1–10	1.34	.52
7F	251	6.34	2.11	1–10	1.35	.55
8M	228	6.75	2.01	2–10	1.32	.52
8F	224	6.88	2.07	1–10	1.29	.57
9M	304	6.94	2.04	2–10	1.28	.57
9F	302	6.88	2.18	1–10	1.26	.63
10M	287	7.45	2.01	2–10	1.20	.60
10F	302	7.30	2.13	2–10	1.19	.65
11M	253	7.63	1.89	2–10	1.16	.58
11F	265	7.09	2.10	2–10	1.22	.62
12M	190	7.55	1.96	1–10	1.18	.60
12F	253	7.32	1.97	2–10	1.20	.58

Table F

Testimony I: Reliability Estimates for the Qualification Subtest

Grade/Sex	N	Mean Score (10 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	6.31	2.01	0–10	1.37	.48
7F	251	6.84	1.75	2–10	1.33	.36
8M	228	6.65	2.02	2–10	1.32	.52
8F	224	7.05	1.97	3–10	1.28	.53
9M	304	6.91	1.98	2–10	1.28	.53
9F	302	7.30	1.90	2–10	1.24	.53
10M	287	7.45	1.92	2–10	1.21	.56
10F	302	7.88	1.73	3–10	1.12	.53
11M	253	7.86	1.67	3–10	1.14	.48
11F	265	7.89	1.72	3–10	1.13	.52
12M	190	7.85	1.91	3–10	1.12	.61
12F	253	7.88	1.72	2–10	1.12	.53

Table G

Testimony I: Reliability Estimates for the Four Reject Subtests as a Unit

Grade/Sex	N	Mean Score (40 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	22.44	5.03	7-37	2.88	0.66
7F	246	22.62	4.83	5-35	2.85	0.64
8M	227	23.37	4.80	12-33	2.85	0.64
8F	224	23.57	5.00	8-35	2.76	0.69
9M	304	23.89	5.08	8-38	2.78	0.69
9F	302	24.09	5.37	7-37	2.69	0.74
10M	287	25.14	4.96	5-36	2.6?	0.71
10F	302	24.99	5.04	7-36	2.54	0.74
11M	253	25.49	4.78	13-35	2.52	0.72
11F	265	24.26	5.50	4-35	2.54	0.78
12M	190	25.49	5.24	10-36	2.56	0.76
12F	252	25.0!	5.31	6-36	2.47	0.78

Table H

Testimony II: Reliability Estimates

Grade/Sex	N	Mean Score (20 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	13.14	3.83	0-20	1.93	.73
7F	251	12.77	3.86	4-20	1.95	.73
8M	228	12.67	4.32	4-20	1.92	.79
8F	224	13.16	4.26	6-20	1.88	.79
9M	304	12.75	4.36	4-20	1.91	.80
9F	302	13.26	4.51	4-20	1.84	.82
10M	287	13.33	4.89	2-20	1.80	.86
10F	302	15.14	4.32	5-20	1.64	.85
11M	253	14.60	4.58	0-20	1.69	.86
11F	265	14.83	4.79	3-20	1.63	.88
12M	190	14.52	4.76	6-20	1.68	.87
12F	253	14.98	4.69	0-20	1.61	.88

Table I

Testimony III: Reliability Estimates for the Total Test

Grade/Sex	N	Mean Score (40 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	23.59	5.58	12–37	2.93	.72
7F	251	23.55	5.33	14–37	2.90	.70
8M	227	24.04	6.14	5–40	2.89	.77
8F	223	25.25	5.29	14–38	2.80	.71
9M	303	24.40	5.99	11–38	2.86	.77
9F	305	25.68	5.72	8–38	2.77	.76
10M	228	26.28	6.54	3–40	2.73	.82
10F	311	27.05	5.83	16–39	2.65	.79
11M	256	27.33	6.66	0–40	2.64	.84
11F	262	28.36	6.07	10–39	2.55	.82
12M	195	26.82	6.42	13–39	2.70	.82
12F	251	27.69	6.33	4–40	2.56	.83

Table J

Testimony III: Reliability Estimates for the Recency Subtest

Grade/Sex	N	Mean Score (20 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	12.57	3.32	4–20	2.00	.62
7F	251	12.88	3.42	7–20	1.97	.65
8M	227	12.73	3.61	6–20	1.97	.69
8F	223	13.73	3.44	7–20	1.86	.69
9M	303	12.91	3.64	5–20	1.95	.70
9F	305	13.90	3.36	7–20	1.85	.68
10M	288	14.15	3.64	7–20	1.83	.73
10F	311	14.69	3.21	6–20	1.75	.71
11M	256	14.54	3.70	0–20	1.76	.76
11F	262	15.40	3.28	7–20	1.66	.73
12M	195	14.30	3.64	6–20	1.79	.75
12F	251	14.97	3.25	7–20	1.69	.72

Table K

Testimony III: Reliability Estimates for the Proximity Subtest

Grade/Sex	N	Mean Score (20 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	11.02	3.22	3–20	2.06	.57
7F	251	10.67	3.07	4–19	2.05	.53
8M	227	11.31	3.53	5–20	2.03	.65
8F	223	11.52	3.09	4–19	2.01	.56
9M	303	11.49	3.40	4–20	2.02	.63
9F	305	11.78	3.42	2–19	1.98	.65
10M	288	12.13	3.87	1–20	1.94	.73
10F	311	12.63	3.71	2–20	1.90	.72
11M	256	12.79	3.86	0–20	1.89	.74
11F	262	12.96	3.81	4–20	1.84	.75
12M	195	12.52	3.64	4–20	1.94	.70
12F	251	12.72	3.99	4–20	1.83	.78

Table L

Reasoning I: Reliability Estimates

Grade/Sex	N	Mean Score (28 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	245	8.24	5.00	0–24	2.18	.80
7F	248	8.97	4.87	0–25	2.24	.78
8M	230	9.41	5.58	1–28	2.23	.83
8F	218	10.79	5.99	3–27	2.24	.85
9M	302	9.69	5.77	1–25	2.23	.84
9F	301	11.61	6.57	1–27	2.24	.88
10M	277	12.90	6.89	1–28	2.25	.89
10F	294	14.48	7.48	2–28	2.16	.91
11M	262	14.08	7.58	1–28	2.18	.91
11F	270	15.21	7.48	1–28	2.16	.91
12M	191	13.86	7.84	1–28	2.16	.92
12F	252	15.70	7.16	3–28	2.18	.90

Table M

Reasoning II: Reliability Estimates

Grade/Sex	N	Mean Score (28 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	246	15.94	3.91	7–26	2.42	.60
7F	251	17.24	4.16	9–27	2.37	.66
8M	227	17.76	4.40	8–28	2.38	.70
8F	223	18.25	4.43	9–27	2.30	.72
9M	303	18.00	4.87	8–28	2.32	.76
9F	305	19.55	4.62	7–28	2.17	.77
10M	288	18.88	4.95	6–27	2.14	.81
10F	311	21.00	4.85	9–28	2.00	.82
11M	256	20.81	5.19	0–28	2.05	.84
11F	262	21.50	4.53	9–28	1.94	.81
12M	195	21.02	5.15	9–28	2.03	.84
12F	251	21.90	4.67	5–28	1.86	.84

Table N

Reasoning III: Reliability Estimates

Grade/Sex	N	Mean Score (28 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	245	12.44	5.67	3–27	2.34	.82
7F	248	13.77	5.24	3–26	2.35	.80
8M	230	13.89	6.13	3–28	2.31	.85
8F	218	15.72	5.89	4–26	2.27	.85
9M	302	14.66	6.36	3–26	2.27	.87
9F	301	16.57	5.92	0–27	2.24	.85
10M	277	17.55	6.22	3–28	2.18	.87
10F	294	19.14	5.42	5–28	2.10	.84
11M	262	18.68	6.25	0–28	2.11	.88
11F	270	19.53	5.63	2–28	2.06	.86
12M	191	18.03	6.71	4–28	2.12	.90
12F	252	20.59	4.94	5–28	2.00	.83

Table O

Reasoning IV: Reliability Estimates

Grade/Sex	N	Mean Score (28 max.)	Standard Deviation	Spread of Scores	Standard Error of Measurement	Hoyt Reliability
7M	245	14.39	4.97	4–27	2.40	.76
7F	248	15.41	4.73	5–27	2.39	.74
8M	230	16.06	4.81	5–25	2.37	.75
8F	218	16.94	4.86	5–27	2.33	.76
9M	302	16.60	5.25	4–27	2.30	.80
9F	301	17.82	5.09	4–28	2.25	.80
10M	277	18.72	5.12	5–28	2.21	.81
10F	294	19.73	4.51	6–28	2.15	.77
11M	262	19.23	4.96	3–28	2.17	.80
11F	270	20.17	4.35	6–28	2.10	.76
12M	191	19.04	5.63	5–28	2.14	.85
12F	252	21.16	3.88	10–28	2.03	.72

GPO 815–294–3

REFERENCES

Allen, R. R., Jerry D. Feezel, and Fred J. Kauffeld. A taxonomy of concepts and critical abilities related to the evaluation of verbal arguments. Occasional Paper Number 9. Wisconsin Research and Development Center, The University of Wisconsin, Madison, August 1967.

Allen, R. R., Jerry D. Feezel, and Fred J. Kauffeld. The Wisconsin tests of testimony and reasoning assessment (WISTTRA). Practical Paper Number 6. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, January 1969.

Allen, R. R., Jerry D. Feezel, Fred J. Kauffeld, and Margaret L. Harris. The development of the Wisconsin tests of testimony and reasoning assessment (WISTTRA). Technical Report Number 80. Wisconsin Research and Development Center, The University of Wisconsin, Madison, April 1969.

Allen, R. R., Thomas Fischbach, and Mary R. Quilling. A field test of the semiprogramed introduction to verbal argument. Wisconsin Research and Development Center. The University of Wisconsin, Madison, 1969 (in progress).

Allen, R. R., Fred J. Kauffeld, and William R. O'Brien. A semiprogramed introduction to verbal argument. Wisconsin Research and Development Center, The University of Wisconsin, Madison, December 1968.

Allen, R. R., and Robert K. Rott. The nature of critical thinking. Theoretical Paper Number 20. Wisconsin Research and Development Center. The University of Wisconsin, Madison, Wisconsin, May 1969.

Harris, Margaret L. A factor analytic study of the Wisconsin tests of testimony and reasoning assessment (WISTTRA). Technical Report Number 100. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, October 1969.

Miller, Donald M., et al. Multivariate procedures for stratifying school districts. Laboratory of Instructional Research, The University of Wisconsin, Madison, 1967.

Thorndike, Edward L. and Irving Lorge. The teacher's wordbook of 30,000 words. Columbia: Teachers College Press, 1944.

FOOTNOTE

[1]The names of the tests are: Testimony I—Appraising Testimony in Terms of Internal Criteria which consists of one Accept subtest and four Reject subtests (Bias, Position, Competence, and Qualification); Testimony II—Appraising Testimony in Terms of External Criteria (Consistency with other Testimony); Testimony III—Appraising Testimony in Terms of External Criteria (Recency and Proximity); Reasoning I—Recognizing and Selecting Warrants in Arguments; Reasoning II—Recognizing Statements which Answer Reservations in Arguments; Reasoning III—Selecting Reservations in Arguments; and Reasoning IV—Selecting Claims in Arguments.

GPO 815-294-2